

**PENDEKATAN ESTIMATOR SPLINE PADA MODEL
REGRESI NONPARAMETRIK DENGAN ERROR
BERDISTRIBUSI LOGNORMAL**

SKRIPSI

MPM 45/04

Tri
P



MILIK
PERPUSTAKAAN
UNIVERSITAS AIRLANGGA
SURABAYA

INDRIYANTI TRISNOWATI

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS AIRLANGGA
SURABAYA**

2004

**PENDEKATAN ESTIMATOR SPLINE PADA MODEL
REGRESI NONPARAMETRIK DENGAN ERROR
BERDISTRIBUSI LOGNORMAL**

SKRIPSI

**Sebagai Salah Satu Syarat Untuk Memperoleh
Gelar Sarjana Sains Bidang Matematika Pada Fakultas Matematika dan
Ilmu Pengetahuan Alam Universitas Airlangga**

Oleh :

INDRIYANTI TRISNOWATI
NIM. 080012125

Tanggal Lulus : 2 Juni 2004

Disetujui Oleh :

Pembimbing I,



Nur Chamidah, S.Si, M.Si
NIP. 132 205 653

Pembimbing II,



Drs. Suliyanto, M.Si
NIP. 131 933 016

LEMBAR PENGESAHAN SKRIPSI

Judul : PENDEKATAN ESTIMATOR SPLINE PADA MODEL
REGRESI NONPARAMETRIK DENGAN ERROR
BERDISTRIBUSI LOGNORMAL

Penyusun : INDRIYANTI TRISNOWATI

NIM : 080012125

Tanggal Ujian : 2 Juni 2004

Disetujui Oleh :

Pembimbing I,

Pembimbing II,



Nur Chamidah, S.Si, M.Si
NIP. 132 205 653



Drs. Sulyanto, M.Si
NIP. 131 933 016

Mengetahui :

Dekan Fakultas MIPA
Universitas Airlangga

Ketua Jurusan Matematika
FMIPA Universitas Airlangga



Drs. H. Abdul. Latief Burhan, MS
NIP. 131 286 709



Drs. Moh. Imam Utoyo, M.Si
NIP. 131 801 397

MILIK
PERPUSTAKAAN
UNIVERSITAS AIRLANGGA
SURABAYA

Indriyanti T., 2004 *Estimator Spline Approach in Regression Nonparametric Model with Lognormal Distribution Error*. This Final Project under guidance of Nur Chamidah S.Si, M.Si and Drs. Suliyanto M.Si, Departement of Mathematics, Faculty of Mathematics and Natural Science Airlangga University

Abstract

The regression analysis is a method in statistica which used to determine relationship between two or more variables. Suppose Y is respon variable and X is predictor variable for n observations, the relationship of that variables can be expressed in a model multiplikatif as follows :

$$y_i = f(x_i) \cdot \varepsilon_i \quad i = 1, 2, \dots, n \quad \varepsilon_i \sim LN(0, \sigma^2)$$

where $f(x_i)$ is unknown regression function and ε_i are observation errors, which assumed to be independence with mean 0 and variance σ^2 . The model would be transformed by adding natural logarithm and the resulting is

$$y_i^* = f^*(x_i) + \varepsilon_i^* \quad i = 1, 2, \dots, n \quad \varepsilon_i^* \sim N(0, \sigma^2)$$

The purpose of this *skripsi* is to estimate regression function, $\hat{f}(x_i)$ with the spline estimator approach. Based on the spline natural estimator approach, we got the estimator form of regression function :

$$\hat{f}(x_i) = e^{T\hat{\alpha} + \Sigma\hat{\beta}} \quad \text{with } \hat{\alpha} = (T' M^{-1} T)^{-1} T' M^{-1} Y^* \\ \hat{\beta} = M^{-1} (Y^* - T\hat{\alpha})$$

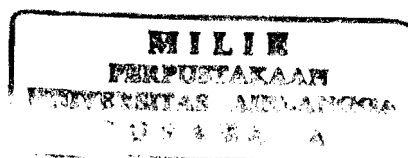
and based on spline least square estimator approach, we got the estimator form of regression function :

$$\hat{f}(x_i) = e^{Z(\lambda)\hat{\theta}} \quad \text{with } \hat{\theta} = [Z(\lambda)' Z(\lambda)]^{-1} Z(\lambda)' Y^*$$

From the result of nonparametric regression model application on the data diare as shown in lampiran 1, which the number of patient diarrhea as variable of respon and percentage of gift ASI as variable predictor, and using spline least squares estimator obtained optimal value lambda $\lambda = \{10, 17, 18\}$ and order polynomial $m = 2$ and

$$\hat{f}(x_i) = e^{-2.348433 + 1.08649 x_i - 1.24792 (x_i - 10)_+ + 3.117862 (x_i - 17)_+ - 3.597234 (x_i - 18)_+}$$

Keywords : Nonparametric Regression, Lognormal Distribution, Spline Estimator



Indriyanti T., 2004 *Pendekatan Estimator Spline pada Model Regresi nonparametrik dengan Error Berdistribusi Lognormal*. Skripsi ini dibawah bimbingan Nur Chamidah S.Si, M.Si dan Drs. Suliyanto M.Si, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam

Abstrak

Analisis regresi merupakan suatu metode statistika yang digunakan untuk menentukan hubungan antara dua variabel atau lebih. Misalkan Y adalah variabel respon dan X adalah variabel prediktor untuk n pengamatan, maka hubungan antara variabel-variabel tersebut dapat dinyatakan dalam model multiplikatif sebagai berikut :

$$y_i = f(x_i) \cdot \varepsilon_i \quad i = 1, 2, \dots, n \quad \varepsilon_i \sim LN(0, \sigma^2)$$

$f(x_i)$ merupakan fungsi regresi yang tidak diketahui, ε_i error random dengan mean 0 dan variansi σ^2 . Model tersebut akan ditransformasi dengan cara mengambil nilai logaritma natural, sehingga dihasilkan model :

$$y_i^* = f^*(x_i) + \varepsilon_i^* \quad i = 1, 2, \dots, n \quad \varepsilon_i^* \sim N(0, \sigma^2)$$

Tujuan dari skripsi ini adalah untuk mengestimasi fungsi regresi $\hat{f}(x_i)$ dengan pendekatan estimator spline. Berdasarkan pendekatan estimator spline alami didapatkan bentuk estimator untuk fungsi regresi adalah

$$\hat{f}(x_i) = e^{T\hat{\alpha} + \Sigma\hat{\beta}} \quad \text{dengan} \quad \hat{\alpha} = (T' M^{-1} T)^{-1} T' M^{-1} Y^* \\ \hat{\beta} = M^{-1} (Y^* - T\hat{\alpha})$$

Dan berdasarkan pendekatan estimator spline kuadrat terkecil didapatkan bentuk estimator untuk fungsi regresi adalah

$$\hat{f}(x_i) = e^{Z(\lambda)\hat{\theta}} \quad \text{dengan} \quad \hat{\theta} = [Z(\lambda)' Z(\lambda)]^{-1} Z(\lambda)' Y^*$$

Dari hasil penerapan model regresi nonparametrik pada data diare pada lampiran 1, dimana banyaknya penderita diare sebagai variabel respon dan persentase pemberian ASI sebagai variabel prediktor maka dengan estimator spline kuadrat terkecil diperoleh nilai lambda optimal $\lambda = \{10, 17, 18\}$ dan orde polinomial $m = 2$ dan

$$\hat{f}(x_i) = e^{-2.348433 + 1.08649 x_i - 1.24792 (x_i - 10)_+ + 3.117862 (x_i - 17)_+ - 3.597234 (x_i - 18)_+}$$

Kata Kunci : Regresi Nonparametrik, Distribusi Lognormal, Estimator Spline